I. Summary[1]

The ZIP file, located at `https://reopeningdata.github.io/` , contains the replication materials for "Reopening the Economy: What Are the Risks, and What Have States Done?" by Enghin Atalay, Shigeru Fujita, Sreyas Mahadevan, Ryan Michaels, and Tal Roded.

The two main do files that need to be run, in order, are `create_dataset.do` and `reopening_figures.do`. The first of the two files constructs the main datasets (`reopening_dataset.dta` and `reopening_dataset_combined.dta`). The second of the two do files produces the figures and tables that are included in the note and accompanying website.

II. On Measuring States' Initial Designation of Essential Industries

The appendix of "Reopening the Economy: What Are the Risks, and What Have States Done?" provides some background on our procedure to construct an industry-by-county dataset of closures (as of April 17, 2020). Here, we review this material, referring specifically to the replication files.

We follow two different procedures, one for a set of states (and counties) with relatively concise closing orders, and a second for states with lengthier orders. For the states with more concise orders, we hand coded the NAICS industries that we identify as closed or open. Our hand-coded industries are collected in `essential_industries_initial.xlsx`. For states with lengthier orders, we develop a list of keywords (collected in `finalkey.csv`) associated with each NAICS industry. Then, we search for these keywords in the text of states' and counties' closing orders. The output of this searching procedure is collected in `initial_other_states_based_on_cisa3.csv`. The code (and raw text from states' orders) to produce `itial_other_states_based_on_cisa3.csv` is given in the folder Create_Keyword_Based_Designations. Using `essential_industries_initial.xlsx` and `initial_other_states_based_on_cisa3.csv` as inputs, the code to produce a STATA database of states' initial closures is given in step 1 of `create_dataset.do`.

III. On Measuring Reopening

We follow a two-step procedure to construct our dataset of reopening paths by industry and county. The first step involves hand collecting information from states' reopening orders, storing this information in excel files. The second step, executed in STATA, involves compiling this information into a structured dataset; in this dataset, observations are NAICS industry-county pairs and variables indicate whether the county-industry pair should be marked as "open" or "closed" on a given week (for each week between May 1 and the most recently available period.) We describe the two steps in turn.

III.A Collecting Information from States' Reopening Orders

[1] Research results and conclusions expressed throughout this project are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or the Federal Reserve Board of Governors.

Our data collection efforts center on reading through states' executive orders and local news stories.

There are two groups of states in this step. First, certain states have clearly defined "phases" and within-state regions, whereby the movement into successive phases varies across regions. For instance, Pennsylvania's reopening plan involves a "red" (initial) phase, a "yellow" phase, and a (terminal) "green" phase. The state's reopening process largely involves each of the 67 counties passing through the three phases. Other states (e.g., New York) have grouped counties into regions that move through the reopening process in tandem.[2] For seven states – Illinois, Michigan, New York, Oregon, Pennsylvania, Virginia, and Washington – we have compiled two worksheets: the movement of counties into different phases, and the list of industries that can be open in each county.

For the second set of states, we instead directly list the NAICS industries that are re-opened on each date. Even for this second group of states, there are certain instances of within-state across-county variation in when industries reopen. For counties that are exceptions to the rest of their respective states, we list these exceptions in our excel file as well. The file that collects information on reopening for these states is `reopening_other.xlsx` . Within this file, there are four worksheets. The first worksheet includes the main information that will be read in by STATA. The second and third worksheets list the counties (and associated FIPS county codes) and the titles of industries (and associated NAICS industry codes). The final worksheet provides a partial list of the sources for this worksheet.

Regarding the mapping between text and NAICS industries, there are three exceptional cases to keep in mind. First, certain states issued blanket reopening orders as part of their reopening. For instance, on June 1 Oklahoma entered "phase 3" of its reopening plan, whereby all industries are allowed to operate (albeit with certain restrictions). In cases like these, we assign all industries to be open. Second, certain states have indicated that "office-based businesses" (or similar phrases) are allowed to reopen. (An example of this would be Rhode Island, in phase 1 of its plan.) In these instances, we take NAICS sectors beginning with "5" (with some exceptions) in addition to NAICS sectors 8132-8139.[3] Third, in their reopening orders, many states mention reopening manufacturing without discussing wholesale and related distributing activities, while other states mention reopening these two sectors in tandem. In our view, it is unlikely that wholesale is still a restricted activity (especially given that manufacturing and retail have been allowed to open in every state.) So, for each state that (i) lists manufacturing as a reopening and (ii) has not, to the current date, explicitly mentioned wholesale as a sector that has re-opened, we designate the re-opening date for wholesale sector (plus distribution support activities and warehousing and storage, NAICS codes 488 and 493 respectively) to be that of the manufacturing sector's reopening date.[4]

---

[2] Within New York, for example, eight counties – Albany, Columbia, Green, Renssalaer, Saratoga, Schenectady, Warren, and Washington – comprise the "Capital District" region. These eight counties moved into phase 1 of New York's plan on May 20, phase 2 on June 3, phase 3 on June 17, and phase 4 on July 1.

[3] The exceptions, within the NAICS 5 sector, are those industries that have been explicitly listed in states' reopening orders. These include 51211 (Motion Picture and Video Production), 51213 (Motion Picture and Video Exhibition), 51912 (Libraries and Archives), 54192 (Photography Services), 54194 (Veterinary Services), and 5617 (Services to Buildings and Dwellings).

[4] Of course, this change only applies to states that had at least partially classified wholesale and relate distribution activities as nonessential in their initial closing orders.

III.B Storing Information in a STATA database

The code to produce a STATA database of states' reopening paths is given in steps 2 to 6 of `create_dataset.do`.

- Step 2 reshapes the excel files corresponding to reopening in Illinois, Michigan, New York, Oregon, Pennsylvania, Virginia, Washington so that it has the same format as all other states.
- Step 3 takes the reshaped files from step 2 and the excel file from other states and stores local macro variables with information on opening dates by counties and industries.
- Step 4 uses the local macro variables from step 3 to construct a database with county-industry pairs as the observation and reopening status (on a weekly basis, from early May to the current period) as different variables. This step also constructs a variable identifying when the industry was closed in its state.
- Step 5 modifies the dataset to account for recent (re-)closures (e.g., Texas' late June closure of bars). The output of this step is `reopening_dataset.dta`
- Step 6 collapses up to a coarser industry definition, the industry definition that allows us to merge to datasets like O*NET and the CPS. The output of this step is `reopening_dataset_combined.dta`

Step 4 takes several hours to run. However, to circumvent running the time-consuming portion of the do file, one may download `reopening_dataset.dta` (from `https://reopeningdata.github.io/reopening_dataset.zip`) and then set the local macro variable skip_step_4_5 equal to 1.

IV. Figures and Tables

We produce the figures and tables in our note and on `https://reopeningdata.github.io/figures.html` using the `reopening_figures.do` do file. In addition to the data on reopening, whose construction we describe above, this do file employs data from the American Time Use Survey (to compute the share of workers who report the ability to work from home in each industry), the County Business Patterns (to compute employment in each industry and county), the Current Population Survey (to compute the age of workers in each industry), O*NET (to compute contact intensity in each occupation), the National Employment Matrix (to compute the number of workers in each industry), and the Quarterly Census of Employment and Wages (to compute the number of workers in each industry). The input files are all collected in in the ZIP file.